Forward Selfies

Philipp Trenz
Hasso Plattner Institute for Digital Engineering, University of Potsdam
Potsdam, Germany
philipp.trenz@hpi.de

Mandy Klingbeil
Digital Masterpieces GmbH
Potsdam, Germany
mandy@digitalmasterpieces.com

Sebastian Pasewaldt
Digital Masterpieces GmbH
Potsdam, Germany
sebastian@digitalmasterpieces.com

Jürgen Döllner
Hasso Plattner Institute for Digital Engineering, University of Potsdam
Potsdam, Germany
juergen.doellner@hpi.de

Matthias Trapp
Hasso Plattner Institute for Digital Engineering, University of Potsdam
Potsdam, Germany
matthias.trapp@hpi.de

(a) Common Selfie (b) Forward Selfie (c) Back Camera Image (d) Front Camera Image and Matte

Figure 1: Comparison of a common selfie captured using the front camera (a) and a forward selfie obtained by our mobile application (b). The forward selfie is obtained by blending the back camera image (c) with the simultaneously acquired front camera image using a person-specific matte (d) that is computed during capturing.

ABSTRACT

Taking selfies is a common practice for smartphone users. Simultaneously capturing oneself and the desired background is not a trivial task, because it is often not possible to get a good view of both. Moreover, users often loose attention of their surroundings, thus taking a selfie also showed to lead to serious injuries. To ease the process of capturing selfies and to make it more safe, this work proposes forward selfies as a simple yet effective concept to account for both, risk and challenges. Forward selfies seamlessly combine images of the front-facing and the rear-facing smartphone camera. We propose a mobile app that builds on this concept and implements the selfie synthesis in a post-processing image composition stage. Thereby, we can take advantage of the commonly more advanced back-camera hardware, i.e., providing higher image resolutions, larger field of views, and different perspectives. Finally, we leverage built-in camera optimizations for independently (de-)focusing objects at different distances, such as for persons and backgrounds. We conclude that the concept of forward selfies can effectively address and solve certain challenges of capturing selfies, which we demonstrate by a simple app user interface.

SIGGRAPH '21 Appy Hour, August 09-13, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8358-5/21/08.
https://doi.org/10.1145/3450415.3464398

CCS CONCEPTS

• **Information systems → Multimedia content creation**; • **Computing methodologies → Computational photography**.

KEYWORDS

mobile devices, forward selfies, mobile camera, portrait matte

ACM Reference Format:
Philipp Trenz, Mandy Klingbeil, Sebastian Pasewaldt, Jürgen Döllner, and Matthias Trapp. 2021. Forward Selfies. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Appy Hour (SIGGRAPH '21 Appy Hour), August 09-13, 2021.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3450415.3464398

1 MOTIVATION

Selfies are digital photos capturing a self portrait. They are typically taken with the front camera of a mobile phone held at arm's length or by using a selfie stick. Acquiring a good selfie includes different challenges during the capturing process [Kalayeh et al. 2015; Yeh and Lin 2014]: C1: Hold the phone and control the camera's orientation; C2: Position oneself in the picture and in relation to a meaningful background scene; C3: Control the head pose and facial expression [Hsieh and Yeh 2017]; C4: Trigger the image capturing without changing C1-C3, and C5: Be alert to the surroundings. This schematized process is often repeated until a sufficient result is obtained, which requires to focus on multiple aspects at once. Thus, taking selfies may lead to serious injuries in some cases, especially at adventurous settings [Virk and Dhall 2019]. To simplify this process and protect users from hazard, we propose the concept of the *forward selfie* that decouples selfie and scene capturing while still enabling high-quality results.

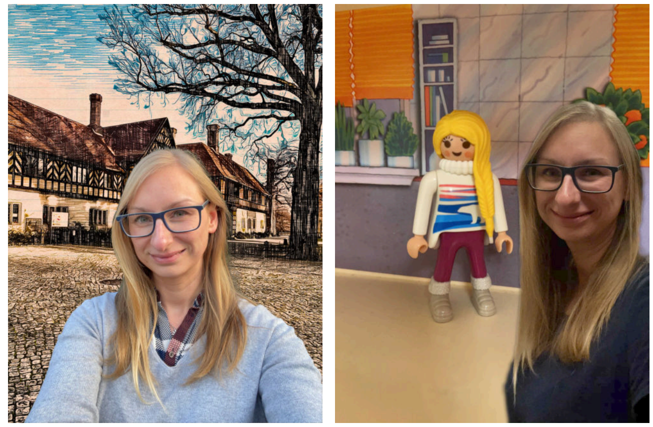
Forward selfies make use of mobile hardware that enables to capture front and back images simultaneously. With our approach, we solved the following challenges: (1) Mattes need to be computed to enable a seamless composition of humans in front of backgrounds, (2) the implemented techniques need to perform in real-time to ensure immediate visual feedback in a preview, and (3) a simple, yet effective user interface should support users during capturing and thus address the challenges described above. Fig. 1 shows a comparison between a traditional selfie and a forward selfie in the same setting captured with our app.

2 APP DESIGN

We implemented the concept of the forward selfie in a mobile app that uses the simultaneous capturing of front and back camera images of modern mobile devices (iPad Pro and iPhone 12). The user is offered two options for taking a selfie: Picture-in-Picture and Preview mode. In the first, the user is presented with the pictures from the front and rear camera superimposed on each other. The latter presents a live preview of the forward selfie with medium segmentation quality. On capture, the images are adjusted and blended based on a portrait matte computed during capturing. In an optional adjustment phase, the user can then realign the selfie on the background image and apply stylization for content-based image manipulation, as shown in Fig. 2.

3 IMPLEMENTATION ASPECTS

We implemented a prototypical application of our concept on iOS with Metal. As test devices, we use an iPad Pro 2020 and an iPhone 12 Pro with TrueDepth™ cameras that support the *Portrait mode* provided by the Apple AVFoundation framework and the simultaneous capturing of photos with multiple cameras. The Portrait mode generates a refined human segmentation matte from RGB and depth information using machine learning—which could also be synthesized with a more generalized but slower tri-map free approach (e.g., *MODNet* [Ke et al. 2020]) if this mode is not available. Compared to the matte generation based on thresholded depth images, as used in the preview view, the portrait matte provides more detail, but takes significant more time to be computed. For example, given an RGB image of 3088×2320 pixels, the corresponding depth map computed with interactive performance only comprises 640×480 pixels, while the generation of a portrait matte with 1544×1160 pixels takes on average 112 milliseconds longer. The mattes can be encoded within regular still images of the HEIC and layered JPG format. This allows a forward selfie not only to be composed from captured photos from within the forward selfie app, but also based on the *Portrait mode*-enabled photographs from the photos library. The adjustment, stylization and compositing techniques are implemented using the Apple CoreImage framework with custom Metal Performance Shaders.



(a) Image stylization (b) Mixed perspectives

Figure 2: Application examples of our forward selfie app. The foreground and background image can be stylized independently, for example to place oneself in a stylized world. Playing with the perspective of both images, yields new forms of selfies.

ACKNOWLEDGMENTS

This work has been partially funded by the German Federal Ministry of Education and Research (BMBF) through grants 01IS18092 ("mdViPro") and 01IS19006 ("KILabor ITSE").

REFERENCES

Yi-Tsung Hsieh and Mei-Chen Yeh. 2017. Head Pose Recommendation for Taking Good Selfies. In *Proc. Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes* (Mountain View, California, USA). ACM, New York, NY, USA, 55–60. https://doi.org/10.1145/3132515.3132518

Mahdi M. Kalayeh, Misrak Seifu, Wesna LaLanne, and Mubarak Shah. 2015. How to Take a Good Selfie?. In *Proc. International Conference on Multimedia* (Brisbane, Australia). ACM, New York, NY, USA, 923–926. https://doi.org/10.1145/2733373.2806365

Zhanghan Ke, Kaican Li, Yurou Zhou, Qiuhua Wu, Xiangyu Mao, Qiong Yan, and Rynson WH Lau. 2020. Is a Green Screen Really Necessary for Real-Time Portrait Matting? arXiv:2011.11961 [cs.CV]

Jitender Singh Virk and Abhinav Dhall. 2019. Garuda: A Deep Learning Based Solution for Capturing Selfies Safely. In *Proc. International Conference on Intelligent User Interfaces: Companion* (Marina del Ray, California). ACM, New York, NY, USA, 43–44. https://doi.org/10.1145/3308557.3308669

Mei-Chen Yeh and Hsiao-Wei Lin. 2014. Virtual Portraitist: Aesthetic Evaluation of Selfies Based on Angle. In *Proc. ACM International Conference on Multimedia* (Orlando, Florida, USA). ACM, New York, NY, USA, 221–224. https://doi.org/10.1145/2647868.2656401